

As a library, NLM provides access to scientific literature. Inclusion in an NLM database does not imply endorsement of, or agreement with, the contents by NLM or the National Institutes of Health.

Learn more: [PMC Disclaimer](#) | [PMC Copyright Notice](#)



Pathogens. 2023 Jan 14;12(1):143. doi: [10.3390/pathogens12010143](https://doi.org/10.3390/pathogens12010143)

In Silico Analysis of a *Drosophila* Parasitoid Venom Peptide Reveals Prevalence of the Cation–Polar–Cation Clip Motif in Knottin Proteins

[Joseph Arguelles](#)^{1,†}, [Jenny Lee](#)^{1,†}, [Lady V Cardenas](#)², [Shubha Govind](#)^{2,3,4}, [Shaneen Singh](#)^{1,3,4,*}

Editor: Stephen A Klotz

[Author information](#) [Article notes](#) [Copyright and License information](#)

PMCID: PMC9865768 PMID: [36678491](#)

Abstract

As generalist parasitoid wasps, *Leptopilina heterotoma* are highly successful on many species of fruit flies of the genus *Drosophila*. The parasitoids produce specialized multi-strategy extracellular vesicle (EV)-like structures in their venom. Proteomic analysis identified several immunity-associated proteins, including the knottin peptide, LhKNOT, containing the structurally conserved inhibitor cysteine knot (ICK) fold, which is present in proteins from diverse taxa. Our structural and docking analysis of LhKNOT's 36-residue core knottin fold revealed that in addition to the knottin motif itself, it also possesses a Cation–Polar–Cation (CPC) clip. The CPC clip motif is thought to facilitate antimicrobial activity in heparin-binding proteins. Surprisingly, a majority of ICKs tested also possess the CPC clip motif, including 75 bona fide plant and arthropod knottin proteins that share high sequence and/or structural similarity with LhKNOT. Like LhKNOT and these other 75 knottin proteins, even the *Drosophila* Drosomycin antifungal peptide, a canonical target gene of the fly's Toll-NF-kappa B immune pathway, contains this CPC clip motif. Together, our results suggest a possible defensive function for the parasitoid LhKNOT. The prevalence of the CPC clip motif, intrinsic to the cysteine knot within the knottin proteins examined here, suggests that the resultant 3D topology is important for their

biochemical functions. The CPC clip is likely a highly conserved structural motif found in many diverse proteins with reported heparin binding capacity, including amyloid proteins. Knottins are targets for therapeutic drug development, and insights into their structure–function relationships will advance novel drug design.

Keywords: host–parasite, *Drosophila*, *Leptopilina*, cysteine knot, knottin fold, Cation–Polar–Cation clip, antimicrobial peptide, heparin-binding motif, miniproteins

1. Introduction

In host–parasite interactions, the parasite must provide offensive pressure while simultaneously keeping the host alive to sustain viable offspring [1,2]. Such conflicting goals have led to the evolution of virulence factors with diverse molecular strategies while maintaining high host specificity. Parasitic Hymenoptera make up more than a million species, are ubiquitous on the planet, and parasitize on a variety of arthropods, mainly insects [3]. Wasps attacking insect hosts have developed an assortment of behavioral and biochemical strategies for success and remain under constant co-evolutionary pressure with their hosts [4,5,6,7].

The *Leptopilina/Drosophila* system is an emerging model for understanding the molecular foundations of anti-wasp responses and wasp virulence strategies [8]. To suppress host immunity, *Leptopilina* females introduce venom factors into their larval hosts during oviposition [4,9,10]. The well-defined innate immune mechanisms in *Drosophila* [11,12,13] provide a valuable context to study the effects of these wasp venom factors. Infection by the specialist *L. bouleari* activates the humoral and cellular immune arms controlled by the fly’s conserved Toll-NF- κ B signaling pathway. The expression of antimicrobial peptide (AMP) genes such as *drosomycin* is activated after *L. bouleari* infection [14]. Toll-NF- κ B signaling also controls blood cell division and development [15,16], steps that control encapsulation and death of wasp eggs [17,18]. In contrast, infection by the generalist *L. heterotoma* suppresses both immune pathways [9,14], and almost all larval blood cells are destroyed by factors in the wasp venom [19,20].

In this study, we analyze the theoretical structure of an *L. heterotoma* peptide, called LhKNOT, identified within organelle-like secretions in the *L. heterotoma* venom but not in *L. bouleari* venom [21]. These immune-suppressive organelles include more than 350 proteins and possess a proteomic profile of eukaryotic microvesicles and a plethora of immunity-related proteins [21,22,23]. LhKNOT adopts the “Inhibitor Cysteine Knot” (knottin) motif [24]. The knottin motif is broadly conserved in a range of taxa, from plants [25,26] to snails [27]. These small peptides (<60 residues) are functionally diverse. Knottins are used offensively, in spider toxins [28] and cone snail venom [27], as well as defensively, e.g., in wound healing activities of cacti [26], and antimicrobial host defense in plants [25,26,29] and invertebrates [30].

The broad range of knottin activity is attributed to the overall structure of its fold, which relies solely on a conserved

motif of six cysteine residues that takes advantage of these cysteines' unique ability to form disulfide bridges. The knottin motif occurs when cysteine residues are observed in a C1-C2-C3-C4-C5-C6 pattern (where “-” represents a number of highly variable residues). Disulfide bridges are then formed in a similarly conserved fashion (C1-C4, C2-C5, and C3-C6), resulting in C1-C4 and C2-C5 forming a macrocycle, while C3-C6 occurs within this macrocycle, resulting in a “knotted” topology. As this knotted topology largely hinges on the conservation of the six cysteine residues, the knottin motif allows for significant variation in the overall sequence [30]. The combination of their small size, high tolerance of sequence variance, and observed structural stability has made knottins a popular scaffold for drug design [30], potentially even for cancer treatments [31].

Various lines of experimental evidence demonstrate or suggest antimicrobial capabilities for different knottins [25,26,29,30,32,33] or gated ion channel interactions [28,34], although the molecular mechanisms by which these functions are carried out are not well understood. Furthermore, it is not known if these antimicrobial functions share a common mechanism of action. Given its discovery from immune-suppressive particles from fly parasites, our goals were to (a) identify LhKNOT-related sequences from diverse organisms, including *L. heterotoma*, *L. boulardi*, and *Ganaspis* parasitoids, that attack *Drosophila* and (b) identify any novel structural motif(s) present in the known or new LhKNOT homologs with the hope that the prediction of such a conserved motif might suggest possible functions for LhKNOT and guide future experimental research.

We show that a theoretical model of LhKNOT possesses the hallmarks of the knottin fold with three antiparallel beta sheets, constrained by disulfide bridges. The LhKNOT model contains a CPC clip, capable of interacting with heparin, a negatively charged glycosaminoglycan (GAG) of significant medical value. LhKNOT also has the potential for interacting with other GAGs, such as keratan sulfate (a GAG that is highly sulfated, similar to heparin) and hyaluronic acid (the single GAG representative that does not require sulfation) [35]. Furthermore, structural homologs of LhKNOT (i.e., bona fide knottin motif proteins) also possess the CPC clip motif that can dock with heparin. Surprisingly, like LhKNOT and these other bona fide knottin proteins, even the *Drosophila* Drosomycin antifungal peptide shares this CPC clip motif in its structure. This motif was also identified in 41 out of 46 additional LhKNOT homologs from *L. heterotoma*, *L. boulardi*, and *Ganaspis* spp. wasps successfully on *D. melanogaster*. These in silico-based observations suggest that the CPC clip motif, inherent in many knottins, may provide the structural basis for modifying immune or virulence functions. We speculate on a possible antimicrobial or virulence-related function for the wasp LhKNOT. The abundance of the knottin fold and its concurrence with the CPC clip motif in most of the newly identified knottin peptides from three *Drosophila* parasitoids suggests an important yet to be determined role for them in parasitoid physiology and/or in the host–parasite interactions.

2. Results

2.1. The LhKNOT Sequence Shows Similarity to Insect and Plant Antimicrobial

Peptides

A BLASTp search of the nr and TSA databases using the full-length LhKNOT sequence (60 amino acids) showed a limited number of significant hits, primarily annotated as “antimicrobial peptides”, from plants and insects. A few fungal sequences were also identified. All results displayed less than 60% amino acid identity to LhKNOT. The top match for LhKNOT in the nr database was an antimicrobial peptide from the common beetle, *Callosobruchus maculatus*, while the top hit in the TSA database was an antimicrobial peptide from the wheat curl mite *Aceria tosichella*. The conservation was most prominent in the region of the only sequence motif associated with LhKNOT, the cysteine knotted ‘antifungal peptide’ signature (pfam11410 domain, $E = 2.8 \times 10^{-3}$). After running PSI-BLAST for detection of remote homologs, additional sequences were revealed. PSI-BLAST showed similar results with common ice plant, *Mesembryanthemum crystallinum*; red mites, *Dinotrombium tinctorium*; parasitoid wasp, *Trichogramma pretiosum*; click beetle, *Ignelater luminosus*; parasitoid wasp, *Trichogramma brassicae*; common pollen beetle, *Brassicogethes aeneus*; ash borer, *Agilus planipennis*; and sawfly, *Neodiprion lecontei*. Using the Hidden Markov Models-driven HMMER webserver as an alternative method of detecting remote homologs, we identified five fungal hits from *Akanthomyces lecanii*, *Cordyceps javanica*, *Cordyceps confragosa*, *Rosellinia necatrix*, and *Beauveria bassiana*.

A search of the PDB for similar sequences retrieved only two significant results, one for a knottin-type antifungal peptide Alo-3 from the insect *Acrocinus longimanus* and the second from the disulfide-constrained wound healing peptide pB1 from *Pereskia bleo*. The details of all close and remote homologs identified using the different sequence similarity approaches are summarized in [Supplementary Table S1](#).

An alignment of LhKNOT to the known structures of typical knottins as well as identified homologs showed the conservation of the six cysteines that form the three disulfide bridges in LhKNOT ([Figure 1A](#)). On its N-terminus, LhKNOT contains a long, helical segment predicted to function as a signal sequence; internally, it contains three beta-sheets, as expected of the secondary structure of knottin peptides ([Figure 1B](#)).

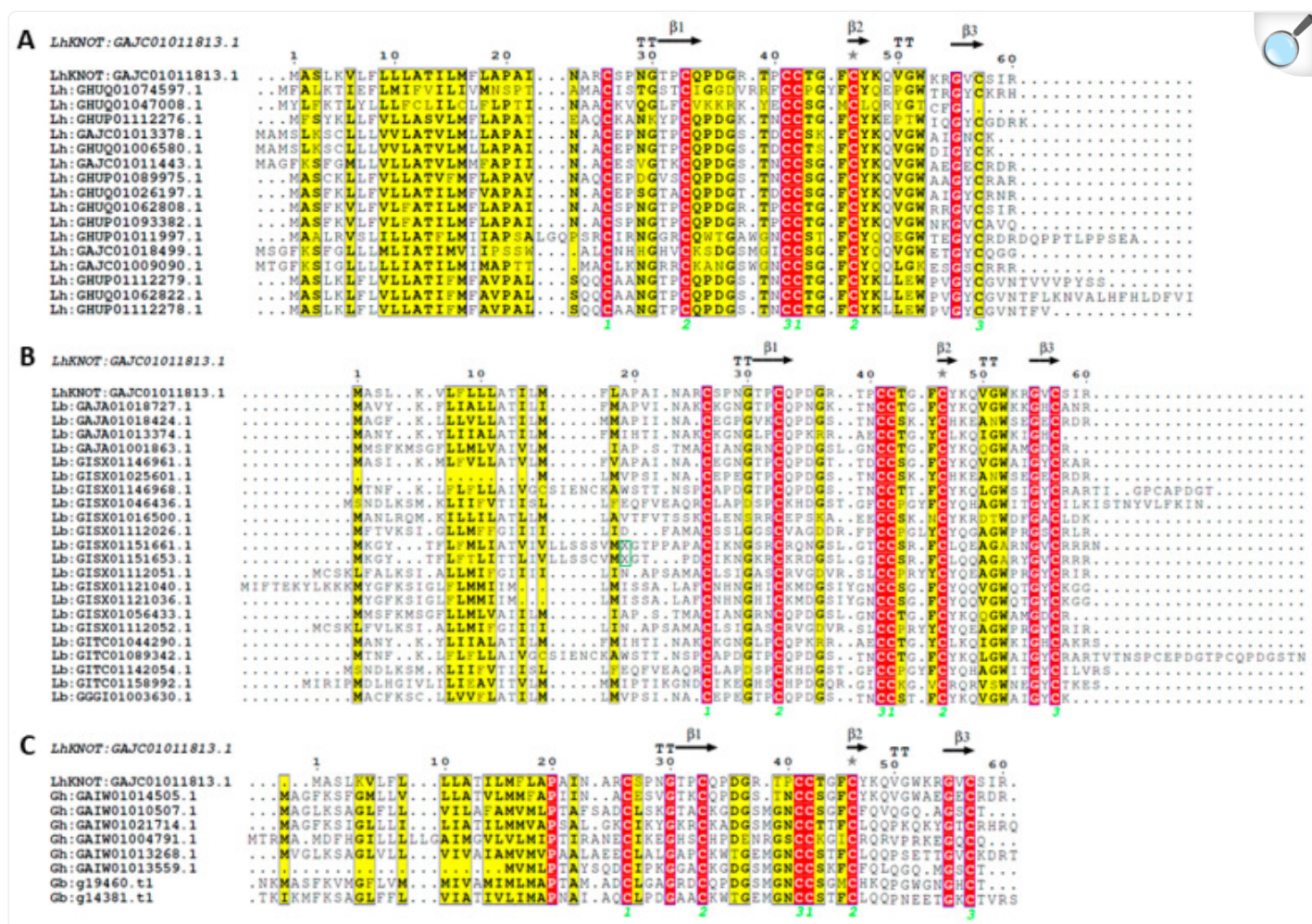
[illegible]

(A) Multiple sequence alignment of LhKNOT (top) and homologs in the knottin family (top-to-bottom): *Mesembryanthemum crystallinum* (plant; antimicrobial; NCBI ID: [AAC19399](#)), *Aceria tosichella* (mite; antimicrobial; NCBI ID:MDE48292.1), *Callosobruchus maculatus* (beetle; unknown; NCBI ID:[VEN35376.1](#)), *Dinothrombium tinctorium* (mite; unknown; NCBI ID:[RWS16713.1](#)), *Trichogramma pretiosum* (Alo-2 like; parasitoid wasp; antimicrobial; NCBI ID: [XP_014233229.1](#)), *Ignelater luminosus*

(beetle; unknown; NCBI ID: [KAF2884324.1](#)), *Trichogramma brassicae* (parasitoid wasp; unknown; NCBI ID: [CAB0031014.1](#)), *Brassicogethes aeneus* (beetle; unknown; NCBI ID: [CAH0563829.1](#)), *Agrilus planipennis* (ash borer; antimicrobial; NCBI ID: [XP_018321119.1](#)), *Neodiprion lecontei* (sawfly; unknown; NCBI ID: [XP_015517007.2](#)), *Akanthomyces lecanii* (fungus; antifungal; Uniprot ID: A0A168C5L6), *Cordyceps javanica* (fungus; antifungal; Uniprot ID: A0A545VVU1), *Cordyceps confragosa* (fungus; unknown; Uniprot ID: A0A179IJT7), *Rosellinia necatrix* (fungus; unknown; Uniprot ID: A0A1S8A723), *Beauveria bassiana* (fungus; unknown; Uniprot ID: A0A0A2VUF6), *Acrocinus longimanus* (Alo-3; beetle; antifungal; PDB ID: 1Q3J), and *Pereskia bleo* (pB1; cactus; PDB ID: 5XBD). Identical residues are highlighted in red, while chemically similar residues are highlighted in yellow. Black arrows denote the conserved cysteine residues that are responsible for the knottin structural fold, and the disulfide bridge connectivity of the three disulfide bridges is marked with neon green numbers. X and XX denote the variable number of residues at the N- and C-terminus, respectively. Secondary structure elements of the modeled LhKNOT are shown at the top. β and T represent β -strand, and β -turn, respectively. **(B)** Secondary structure prediction for LhKNOT, showing a long helical segment (pink cylinder; signal peptide) followed by three beta strands (yellow arrows). This secondary structure follows the canonical knottin signature [36]. **(C)** A prototype of the knottin structural fold from the KNOTTIN database [36] (permission to reproduce this image given by Dr. Jean-Christophe Gelly, INSERM, October 2019). The two disulfide bridges (green) form a macro-cycle, while a third (blue) pulls through the center giving the protein its namesake knotted topology. **(D)** Predicted tertiary structure of LhKNOT, showing the knotted topology and conserved cysteine motif, cysteine residues form disulfide bridges in the conserved pattern seen throughout the knottin family (C_1 - C_4 and C_2 - C_5 form the macro-cycle, C_3 - C_6 pulls through the center).

Additionally, an assessment of sequence similarity of LhKNOT targeted to the available *L. heterotoma* and *L. bouhardi* transcripts and from another *Drosophila* parasitoid, *Ganaspis* spp. (see Methods), revealed an abundance of related sequences. For *L. heterotoma* (*Lh14*), we identified 16 additional *Lh14* knottin proteins ([Supplementary Table S2](#)). For *L. bouhardi*, we identified 22 knottin proteins ([Supplementary Table S3](#)). Six LhKNOT homologs were similarly identified from the *Ganaspis hookeri* (G1) and two LhKNOT homologs were found in the protein predictions made from the *G. brasiliensis* (VA) genome ([Supplementary Table S4](#)). [Figure 2](#) shows the conservation of these parasitoid knottin sequences with LhKNOT. Thus, at least the few *Drosophila* parasitoids sampled in this study appear to encode multiple knottin proteins. Additional such knottin peptides may be encoded by these wasps, and systematic proteomic/bioinformatic approaches are needed to obtain a comprehensive list. The functions of these putative knottins in parasite physiology remain unknown.

Figure 2.



[Open in a new tab](#)

(A) Multiple sequence alignment of LhKNOT (top) and homologs in *Lh14*. (B) Multiple sequence alignment of LhKNOT (top) and homologs in *L. boucardi*. The X marked in the green box corresponds to the following insertions in two *Lb17* sequences: Lb:GISX01151661.1 X = SQAPPPTPEPFHPPGTPP; Lb:GISX01151653.1 X = SQLTSPRSTFPPTGSTIPSIIIT. (C) Multiple sequence alignment of LhKNOT (top) and homologs in *Ganaspis* spp. Identical residues are highlighted in red, while chemically similar residues are highlighted in yellow. Black arrows denote the conserved cysteine residues that are responsible for the knottin structural fold, and the disulfide bridge connectivity of the three disulfide bridges is marked with neon green numbers. Secondary structure elements of the modeled LhKNOT are shown at the top. β and T represent β -strand, and β -turn, respectively. Grey stars added on the top of blocks of sequences, above residues represent residues modeled with alternate conformations.

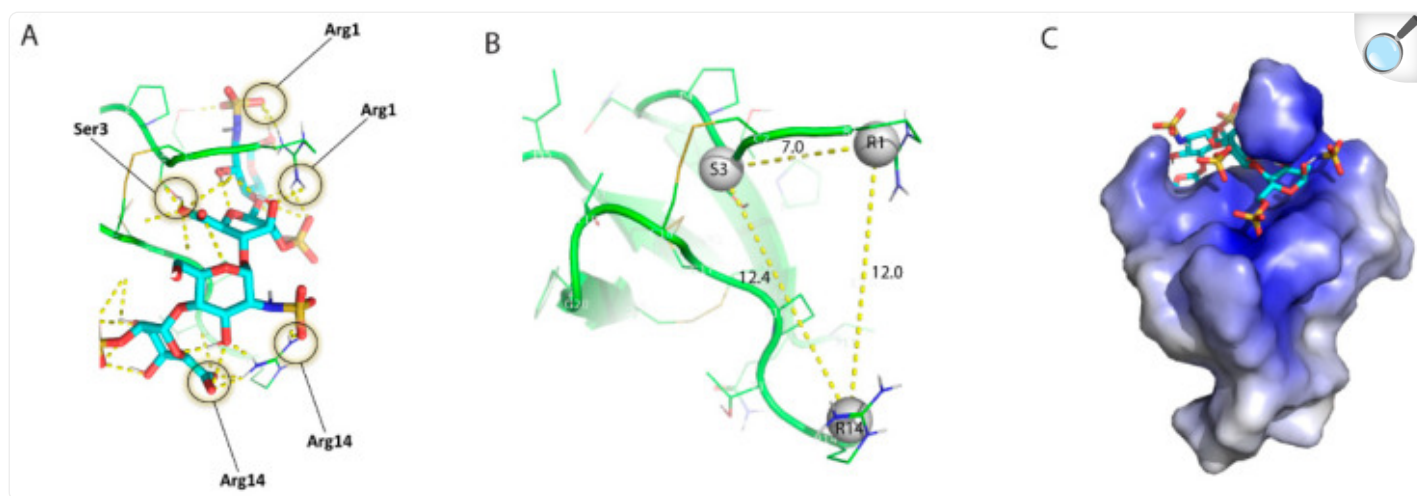
2.2. The Three-Dimensional Structure of LhKNOT

The typical structural fold of a knottin peptide is defined by six cysteines involved in three disulfide bonds to form a cysteine-knotted three-stranded antiparallel beta-sheet [25]. An important characteristic of this fold is the connectivity of the cysteine residues to form the disulfide bridges and the disulfide bridge of cysteines 3 and 6 going through disulfides 1–4 and 2–5 to make a special disulfide through disulfide knot (Figure 1A,C). This disulfide, through a disulfide knot feature, distinguishes the knottin fold from other cysteine knot proteins. Other than the conserved cysteines, the rest of the amino acid sequence can be quite variable, and therefore, only a limited number of knottins are identified by sequence similarity approaches for LhKNOT (see the previous section and [Supplementary Table S1](#)). Fold recognition algorithms and template-based modeling approaches, on the other hand, retrieved several diverse knottins as suitable candidates for modeling LhKNOT. Among the various LhKNOT models generated, the best model was a multi-template-based model of LhKNOT created using the program HHpred [37] with structural restraints derived from the top three templates, Alo-3 (PDB: 1Q3J) [30], omega-agatoxin-IVA (PDB: 1IVA) [28], and PAFP-S (PDB: 1DKC) [25]. Knowledge-based energy profiles calculated using ProSA-web [38] for this selected model showed low energies and a z-score of –4.53, which is comparable to solved structures. Verify3D [39] assessed this multi-template-based model with a passing score (100.00% of the residues in the model have an averaged 3D-1D score greater than or equal to 0.2). The VoromQA [40] score of 0.325 is likely a reflection of the small size of the peptide and the scores of the templates were in the same range ([Supplementary Figure S1](#)). This multi-template-based model typifies the knottin fold with three antiparallel beta strands, constrained by disulfide bridges and the expected connectivity of disulfide bridges (Figure 1D). KNOTER3D, a tool in the KNOTTIN Database [36] corroborated that this model has a bona fide knottin fold and confirmed the previously identified disulfide bonding pattern in [Figure 1D](#).

2.3. LhKNOT Shows the Presence of a CPC Clip, Capable of Interacting with Heparin and Other GAGs

To search for potential LhKNOT functions, structure–function correlations of close structural homologs of LhKNOT were investigated. A knottin protein from the cactus *Pereskia bleo* (PDB: 5XBD) [26] possesses a heparin-binding motif known as the CPC clip. This structural motif forms a geometrically constrained clip of two cationic and one polar residues that interact with heparin by salt bridges and hydrogen bonds clamping it in position within the positively charged surface of the clip [41]. A structural superposition of the cactus knottin and LhKNOT showed that the two structures are highly similar, with an overall difference of less than 2 Å. Docking analysis of LhKNOT with heparin, using ClusPro's specific parameters for heparin as a ligand, revealed that for LhKNOT, the polar interactions fitting the CPC clip motif occur in residues R1-S3-R14 (Figure 3A). Measured distances (both between alpha-carbons and centers of mass of participating residue side chains) differed from those of the CPC clip by less than 1 Å, suggesting the presence of a putative CPC clip in LhKNOT (Figure 3B). The surface electrostatics of the modeled LhKNOT also showed the presence of the expected cationic binding pocket formed by the CPC clip [41,42] (Figure 3C).

Figure 3.



[Open in a new tab](#)

(A) Visualization of the CPC clip motif in LhKNOT. Hydrogen bonds between heparin (teal) and participating residues (Arg1, Ser3, Arg14) of LhKNOT (green) are indicated by dashed yellow lines. (B) Distances measured in PyMOL between the centers of mass of participating residues (grey pseudo-atoms) in the LhKNOT-heparin interaction. Distances are marked on the yellow dashed lines and fall within the expected value ranges of the CPC clip motif. (C) Surface electrostatics of LhKNOT, in complex with heparin. The positively charged surface formed by the cationic residues involved in the CPC clip allows for electrostatic interactions with the overall negatively charged heparin molecule. The surface electrostatic potentials are color-graded from -4 kT/e (red) to $+4$ kT/e (blue).

To assess the CPC clip's ability to interact with other GAGs, additional docking analysis was performed with LhKNOT and keratan sulfate (PDB ID: 1KES) and with LhKNOT and hyaluronic acid (PDB ID: 1HYA). In both cases, the positively charged binding surface created by the CPC clip bound to the negatively charged molecules. Ligplots generated by PDBSum confirmed an association of the GAGs with the CPC-forming residues ([Supplementary Figures S2 and S3](#)).

2.4. Other Structural Homologs of LhKNOT Also Contain the CPC Clip Motif

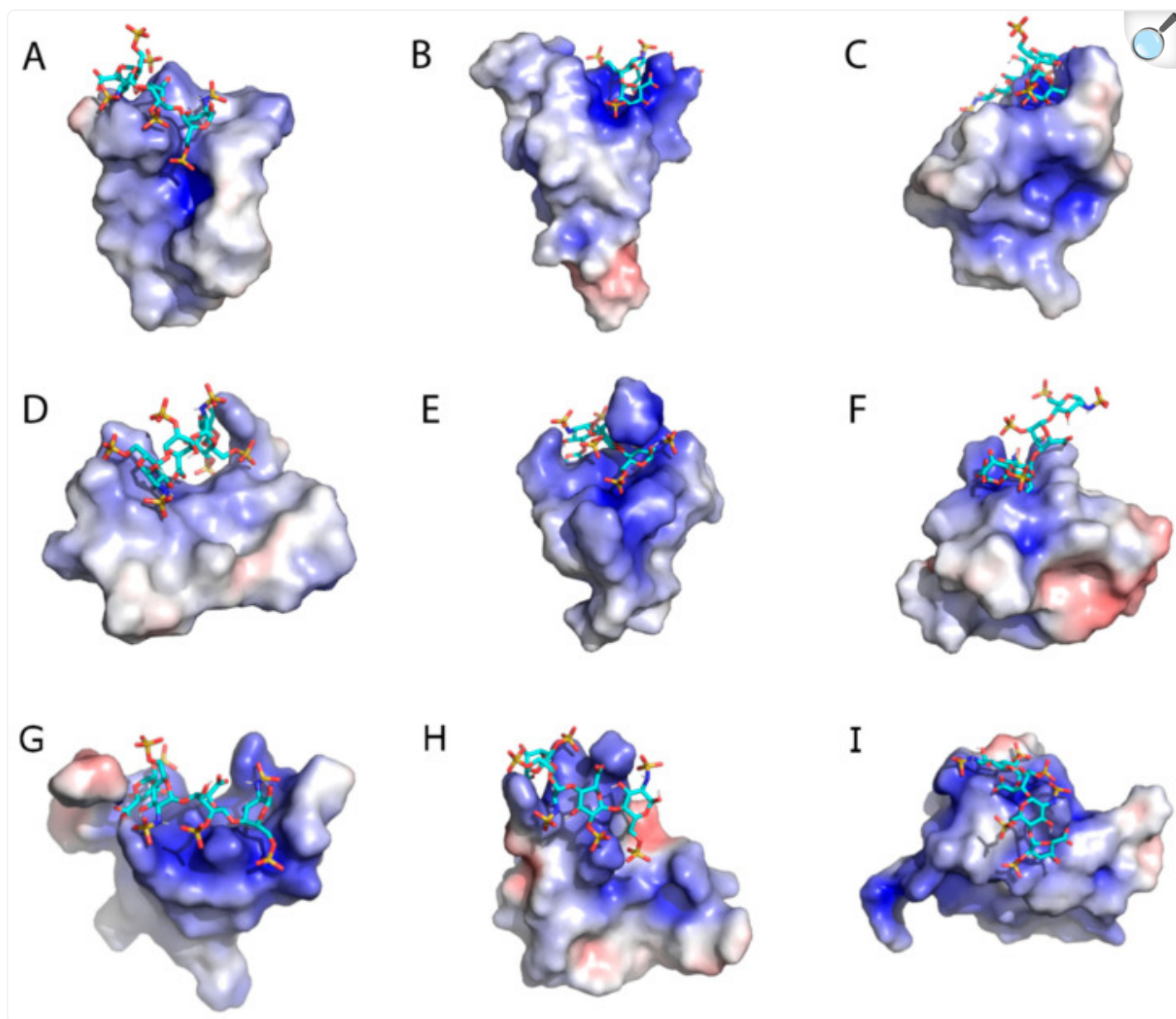
Presence of the CPC clip in both LhKNOT and its structural homolog from *Pereskia bleo* (pB1) suggested the possibility of a broader conservation of the CPC clip within the knottin family. An analysis of 21 close structural

homologs of LhKNOT with experimentally solved PDBs in the KNOTTIN database revealed that, in every case, the peptides possess the CPC clip motif. For most peptides (e.g., *T. tridentatus*), the motif parameters deviated less than 1 Å from the core knottin fold, while for others (e.g., the offensive knottins), the deviation was up to a margin of 3 Å ([Supplementary Table S5](#)). Similarly, modeled knottin proteins of the identified homologs showed that almost all of them possessed the CPC clip.

Sixteen additional sequences from *L. heterotoma*, 22 sequences from *L. bouleari*, and 8 sequences from *Ganaspis* spp., with no known 3D structures, were modeled and analyzed for the presence of the CPC clip motif. While 13 of the 16 putative *L. heterotoma* knottins and 20 of 22 putative *L. bouleari* knottins revealed a CPC clip motif, all *Ganaspis* sequences contained the CPC clip ([Supplementary Table S5](#)). Interestingly, the seven homologs that did not exhibit a CPC clip motif belong to (a) the common pollen beetle sequence (*B. aeneus* ([CAH0563829.1](#)), (b) fungal sequence *R. necatrix* (A0A1S8A723), and (c) three *L. heterotoma* knottins 4,5,6 and *L. bouleari* knottins 3 and 20 ([Supplementary Table S5](#)).

A survey of knottins in the KNOTTIN database, unrelated to LhKNOT, but representative of various taxa (cone snail, horseshoe crab, insect, plant, scorpion, and spider), showed the CPC clip motif. In all cases, the surface electrostatic profile of the CPC clip motif is positively charged, likely, to facilitate interactions with the negatively charged heparin ([Figure 4](#)). However, similar to the beetle, fungal, and wasp knottins described above, the CPC clip was not identified in knottins from the plant cyclotide and sponges ([Supplementary Table S5](#)).

Figure 4.



[Open in a new tab](#)

Heparin-docking results and surface electrostatics of (A) antifungal peptide of *Phytolacca americana* (PFAP-s; plant; PDB: 1DKC), (B) calcium channel selective omega-agatoxin of *Agelenopsis aperta* (Omega-agatoxin IV-A; spider; PDB: 1IVA), (C) antifungal peptide of *Acrocinus longimanus* (Alo-3; beetle; PDB: 1Q3J), (D) wound-healing peptide of *Pereskia bleo* (pB1; cactus; PDB: 5XBD), (E) LhKNOT, (F) antifungal peptide “Drosomycin” of *Drosophila* spp. (Drosomycin; fly; PDB: 1MYN), (G) Conotoxin GS (Conotoxin GS; cone snails; PDB: 1AG7), (H) antimicrobial peptide Tachystatin A isolated from Horseshoe crabs (Tachystatin A; horseshoe crabs; PDB: 1CIX) and (I) Chlorotoxin, a small scorpion toxin of *Leiurus quinquestriatus*

(Chlorotoxin; scorpion; PDB:1CHL) showing the cationic CPC clip motif. Figures oriented to best display the CPC clip structure. The surface electrostatic potentials are color-graded from -4 kT/e (red) to $+4$ kT/e (blue).

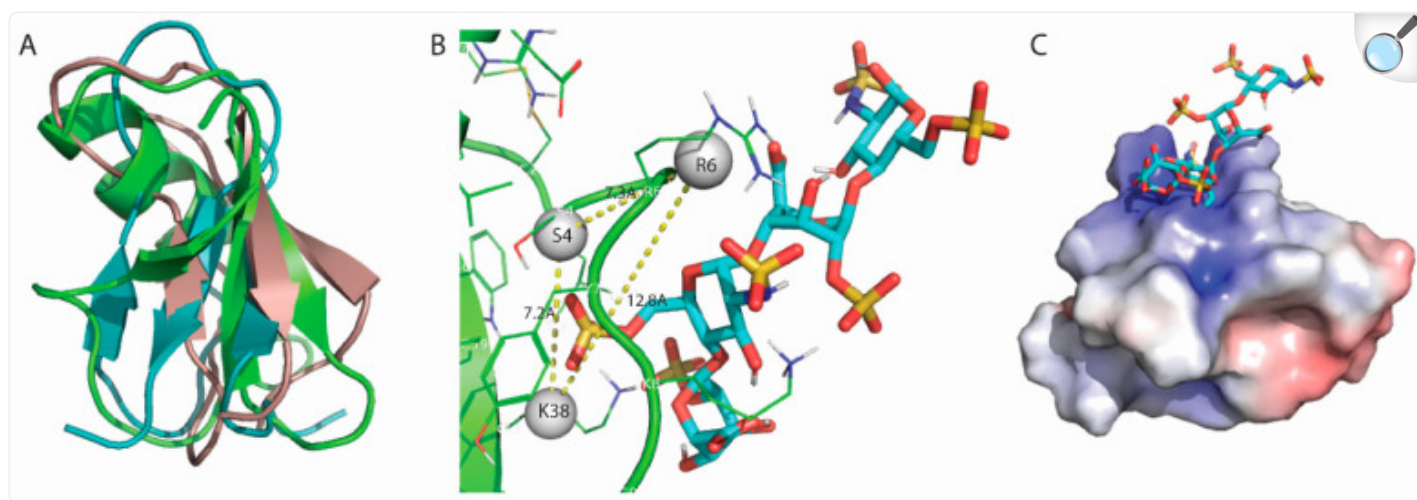
2.5. In Silico Mutational Analysis of LhKNOT CPC Clip Reveals a Secondary CPC Clip

Models of three in silico LhKNOT mutants (R1A, S3A, R14A) were investigated to compare binding interactions of wild type and mutant proteins. In the case of the R14A mutant, heparin was shown to surprisingly associate with a secondary CPC clip involving the R1, S3, R30 residues instead of the R1, S3, R14 motif in the wild type ([Supplementary Figure S4](#)). Double or triple mutants (R14A-R1A; R14A-S3A-R1A), however, did not show an interaction with any of the primary or secondary CPC clip motif residues ([Supplementary Figure S4C,D](#)). The in silico mutants with the R1A mutation (interrupting both the primary and secondary CPC clip motifs), however, did show heparin associating with R29 and R30, as they provide the only remaining cationic surface ([Supplementary Figure S4C,D](#)). Interestingly, the ligand prefers to associate with the CPC clip motif, both in the case of the primary and secondary clips, over this additional cationic surface when the motif is available, suggesting the specific interaction of the CPC clip may allow a more stable association over a non-specific electrostatic interaction.

2.6. A CPC Clip Is Also Present in the Antimicrobial Peptide Drosomycin

Unlike LhKNOT, the structure of Drosomycin contains a different fold with an alpha-helix and a twisted three-stranded beta-sheet, which is stabilized by three disulfide bridges [43]. This structural motif, termed a “cysteine stabilized $\alpha\beta$ -motif”, is also found in other host-defense proteins such as the antibacterial insect defensin A, scorpion toxins, plant thionins, and antifungal plant defensins [43,44]. Given the known antimicrobial function of Drosomycin and its related “cysteine stabilized $\alpha\beta$ -motif” [43,44], we examined the possibility of a common basis for antimicrobial activity and shared structure–function relationships with LhKNOT to better understand the potential significance of this motif in LhKNOT in the context of host defense. Despite a slightly different structural fold, Drosomycin overlays within an RMSD of 2.2\AA of both LhKNOT and pB1 ([Figure 5A](#)). When the previous CPC clip analysis was extended to Drosomycin, the results revealed the presence of the CPC clip motif in Drosomycin as well. The CPC clip motif in Drosomycin (R6-S4-K38) falls within a 2\AA deviation of the motif parameters ([Figure 5B](#), [Supplementary Table S5](#)). Additionally, surface electrostatics of Drosomycin confirmed the putative CPC clip’s formation of a cationic binding pocket as also observed for LhKNOT and other knottins ([Figure 5C](#), [Supplementary Table S5](#)). These results suggest that there may be similarity in the mode of action of Drosomycin and LhKNOT and that the biochemistry embedded in this structural motif may drive antimicrobial function in both peptides.

Figure 5.



[Open in a new tab](#)

(A) Superposition of LhKNOT (blue), wound-healing peptide of *Pereskia bleo* (tan; pB1;cactus; PDB: 5XBD) and Drosomycin (green; Drosomycin; fly; PDB: 1MYN). (B) Distances measured in PyMOL between the centers of mass of participating residues (grey pseudo-atoms) in the Drosomycin–heparin interaction. Distances fall within 2Å of the value ranges of the CPC clip motif. (C) Surface electrostatics of Drosomycin (fly; PDB: 1MYN), showing the positively charged binding surface formed by the CPC clip where heparin is docked. The surface electrostatic potentials are color-graded from -4 kT/e (red) to $+4$ kT/e (blue).

3. Discussion and Speculation

Biochemical functions of proteins are facilitated by their overall structures, which are reportedly 3–10 times more well-conserved than their primary structures [45]. Consequently, sequence alignments often fall short in investigating the functional potential of proteins such as knottins, where the varied primary structures may obscure shared structure–function relationships. In silico methods facilitate broad analysis of potential structure–function correlations for such interesting folds [46,47,48,49,50]. Our detailed analysis of the LhKNOT structure and its structural homologs supports the idea that this parasite protein has the hallmarks of a knottin peptide with functions in host defense: (a) the closest matches to LhKNOT are antimicrobial peptides; (b) LhKNOT’s Pfam signature describes it as cysteine knotted “antifungal peptide”; and (c) a CPC clip is present.

Analysis of a knottin wound-healing peptide bleogen pB1 from the leafy cactus *Pereskia bleo* initially revealed the

presence of a CPC clip, a conserved structural signature of heparin-binding proteins [42]. It has immunity- and signaling-related functions in wound healing and cancer and is used as an anticoagulant [51]. GAGs are present on animal cell surfaces and have become an increasingly popular target of studies ranging from clinical to cosmetic [52]. The CPC clip motif of heparin-binding proteins is considered to be the minimum required motif for heparin-binding affinity and is defined by specific spatial relationships between three residues, two cationic (Arg/Lys) and one polar (Asn, Gln, Ser, Thr, Tyr). This CPC clip motif is defined by distances between both the α carbons of the involved residues as well as the distances between the centers of mass of the involved residues. Heparin-binding proteins are known to interact promiscuously with other negatively charged substrates, most notably the surface membranes of microbes [41]. As a result, heparin-binding proteins can exhibit antimicrobial activity to varying degrees [53]. While the precise mechanisms of this antimicrobial potential are not well defined, it has been hypothesized that the CPC clip may facilitate these protein–microbe interactions [41]. New potential roles of heparin binding are emerging, such as in the potential therapeutic interventions for amyloidogenesis [54], so understanding the structural determinants of the CPC clip in diverse proteins may be key to delineating the mechanism of the interaction and its disruption.

A clue to a possible antimicrobial mechanism for LhKNOT comes from our finding that its inhibitory cysteine knot and those of other knottins examined here have a CPC clip intrinsic to their folds ([Supplementary Table S5](#)). It has been noted that the structural and biophysical requirements for heparin-binding proteins' ability to interact with heparin are strikingly similar to the structural features of many known AMPs [53]. Thus, the CPC clip motif may fit the criteria for both these requirements [41]. Our discovery of this CPC clip motif not only in LhKNOT but also in 75 other related or unrelated knottin proteins ([Supplementary Table S5](#)) suggests that this CPC clip motif may be a highly conserved feature of the knottin proteins. The coincident occurrence of the CPC clip motif within the cysteine knot structure in diverse proteins suggests that the CPC clip motif itself may underlie the antimicrobial function. However, because this structural motif is also identified in offensive knottins ([Supplementary Table S5](#)), it may also serve a more general role via this more generalized CPC clip interaction motif.

The discovery of a putative CPC clip motif in Drosomycin is intriguing, as it sheds light on the possible mechanism of action of this antifungal peptide and supports a possible antimicrobial function for the CPC clip motif itself. Previous studies have shown the importance of Drosomycin's residues R6 and K38, which are part of its CPC clip motif (R6-S4-K38), for its antifungal potential; alterations of these residues reduce antifungal activity [55]. Our structural analysis of the wasp LhKNOT and its similarities with *Drosophila* Drosomycin offers possible insights into the *L. heterotoma*'s ability to broadly suppress Toll-NF- κ B pathway in *D. melanogaster* and yet succeed on immune-compromised hosts. In the absence of a functional Toll immune pathway, and consequently without Drosomycin production, *Drosophila* larvae can often succumb to opportunistic pathogens [11,13,56]. LhKNOT's putative antimicrobial activities might defend the host from opportunistic microbial pathogens in wasp-infected, immune-compromised host larvae and developing pupae. Alternatively, similar to Drosomycin, LhKNOT could induce hemocyte apoptosis [57]. *L. heterotoma* EVs enter hemocytes after infection [58], and their activities within EVs compromise encapsulation [19,21].

The CPC clip has been shown to associate promiscuously with GAGs [41], which are nevertheless ubiquitous in insects including *Drosophila* [59]. Heparin sulfate GAGs in flies serve as growth factor receptors and participate in creating and maintaining morphogenic gradients [60]. The CPC clip has also been shown to associate promiscuously with lipopolysaccharides (LPS) [41]. Bacterial LPS, being major surface components of Gram-negative bacteria, are extremely potent stimulators of the innate immune response in various eukaryotic species including insects [61]. GAGs in *Drosophila* also control the binding of α C protein, a virulence determinant of group B streptococcus [62]. Thus, it is possible that a CPC clip containing LhKNOT and Drosomycin (and possibly other fly AMPs, such as Defensin [63]) can modulate immune signaling (and/or binding to the pathogen's virulence proteins), similarly affecting the bacterial pathogen/metazoan parasite outcome. Previous studies have shown that three homologous knottins (Alo-1, 2, 3) from the harlequin beetle *Acrocinus longimanus* are active against yeast, *Candida glabrata* [30]. Additionally, a pore-forming candidacidal peptide, *Psacothasin*, attacks *Candida albicans* [64]. Lastly, a plant pathogenic rust fungus, *Melampsora larici-populina*, produces a knottin peptide as its own candidate effector [65].

Our finding that multiple putative knottin peptides may be secreted by adult parasitoid wasps of *Drosophila* was unexpected and intriguing. While their functions are not known, the possible existence of a knottin gene family suggests some redundancy in their functions that may be linked to the parasitic life history. The exploration of the knottin structural fold in unrelated and yet-to-be discovered knottins will prove to be valuable. Our work provides a short list of residues significant to the proposed physiological activity for knottin family protein investigation. Such information can be incorporated into rational drug discovery and design for infectious diseases and related therapies.

4. Methods

4.1. Sequence Analysis

Initial scans of the NCBI non-redundant protein sequence (nr) database (<https://www.ncbi.nlm.nih.gov/protein>; accessed on 4 April 2018), the Transcriptome Shotgun Assembly (TSA) sequences database (<https://www.ncbi.nlm.nih.gov/genbank/tsa>; accessed on 16 April 2018) [66], and the Protein Data Bank (PDB; <https://www.rcsb.org> ; accessed on 5 July 2018) [67] with the full-length amino acid sequence of the *L. heterotoma* knottin peptide (GAJC01011813.1) using NCBI BLAST (BLASTp; <https://blast.ncbi.nlm.nih.gov/Blast.cgi>; accessed on 16 April 2018) [68] were used to retrieve related protein sequences in other organisms. Remote homologs were searched using Position-Specific Iterative BLAST (PSI-BLAST; <https://blast.ncbi.nlm.nih.gov/Blast.cgi>; accessed on 4 April 2022) [69], and Hidden Markov Models (HMMER; <https://www.ebi.ac.uk/Tools/hmmer> ; accessed on 6 April 2022) [70,71]. BLAST searches were adjusted to word size 2, and scoring matrix BLOSUM45, while default parameters were used for HMMER.

To identify knottin homologs in select *Leptopilina* and *Ganaspis* spp. not accessible in the protein databases, targeted searches using translated BLASTs against genomic and transcriptomic datasets were employed. To identify additional *L.*

heterotoma knottins, the *L. heterotoma* female abdominal transcriptome GAJC [72] and the whole-body transcriptomes from Ground GHUQ (GenBank Accession: [GHUP000000000](#); TSA: *Leptopilina heterotoma* strain Lh14, transcriptome shotgun) and Space GHUP (GenBank Accession: [GHUQ000000000](#); TSA: *Leptopilina heterotoma* strain Lh14, transcriptome shotgun) were searched using tblastn (e-value threshold set to 1, BLOSUM62, low complexity filter enabled, query coverage filter set to 20 percent) against LhKNOT. The following transcriptomes were similarly searched to identify *L. boucardi* knottins: *L. boucardi* female abdominal transcriptome GAJA [72] and the whole-body transcriptomes from Ground GIRC (GenBank Accession: [GIRC000000000](#); TSA: *Leptopilina boucardi* strain Lb17, transcriptome shotgun), Space GISX (GenBank Accession: [GISX000000000](#); TSA: *Leptopilina boucardi* strain Lb17, transcriptome shotgun), and [GGGI000000000](#) (female abdomen/head) [73]. In addition, targeted searches for *Ganaspis* spp. homologs were carried out. For *G. hookeri*, the [GAIW000000000](#) TSA: *Ganaspis* spp. G1 dataset [74] was searched (tblastn, e-value threshold set to 1, BLOSUM45, low complexity filter enabled, query coverage filter set to 20 percent) using the LhKNOT query. The Expasy tool (web.expasy.org/translate/; accessed on 9 December 2022) was used to translate these TSA sequences [75]. The resulting predicted peptide sequences were manually curated based on the presence of the cysteine motif. For *G. brasiliensis*, AUGUSTUS-based (<http://bioinf.uni-greifswald.de/augustus> ; accessed on 9 April 2020) [76] gene predictions from the partially assembled genomic sequence (GenBank:GCA_009823575.1) were made using the *Nasonia vitripennis* genome training module. The resulting predicted proteins were searched using blastp (e-value threshold set to 1, low complexity filter enabled, query coverage filter set to 20 percent) against LhKNOT. Alignments and visualization were performed using EBI's Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo> ; accessed on 9 December 2022) [77] to confirm the presence of the knottin motif and for further curation of the sequences. The set of identified sequences were aligned to LhKNOT to create a multiple sequence alignment using T-COFFEE (<https://www.ebi.ac.uk/Tools/msa/tcoffee> ; accessed on 9 December 2022) [78] and visualized using ESPript3 (<https://esprict.ibcp.fr/ESPript/ESPript/index.php> ; accessed on 9 December 2022) [79] to identify any conserved residues and/or motif(s) in the aligned sequences.

4.2. LhKNOT Primary Structure Analysis

The LhKNOT sequence was also analyzed by the domain/motif detection database Conserved Domain Database (CDD; <https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>; accessed on 6 April 2018) and Pfam (pfam.xfam.org; accessed on 6 April 2018) to identify any domains or sequence signatures [80,81]. The sequence was further analyzed using the signal sequence prediction software SignalP-5.0 (<https://services.healthtech.dtu.dk/service.php?SignalP-5.0> ; accessed on 6 April 2018) [82] and Phobius (<https://www.ebi.ac.uk/Tools/pfa/phobius> ; accessed on 6 April 2018) [83] to identify the location of the signal sequence and allow for a more accurate characterization of the peptide. The mature peptide of 36 amino acids (with the predicted signal sequence, residues 1–24, removed) was used for further analysis. The secondary structure of the peptide was predicted using the meta-server Sympred (<https://www.ibi.vu.nl/programs/sympredwww> ; accessed on 6 April 2018) [84] and analyzed to confirm its correspondence to the expected secondary structure for a knottin fold: an initial alpha-helical segment (predicted signal sequence) followed by three anti-parallel beta strands.

4.3. Tertiary Structure Prediction and Evaluation of LhKNOT and Its Homologs

The programs I-Tasser (<https://zhanggroup.org/I-TASSER> ; accessed on 16 April 2018) [85], Modeller (<https://salilab.org/modeller> ; accessed on 16 April 2018) [86], and HHPred (<https://toolkit.tuebingen.mpg.de/tools/hhpred> ; accessed on 6 April 2018) [37] were used to model the tertiary structure of LhKNOT. Three structural homologs with the knottin-like fold Alo-3 (PDB: 1Q3J) [30], omega-agatoxin-IVA (PDB: 1IVA) [28], and PAFP-S (PDB: 1DKC) [25] were identified as top candidates and used as templates in model building programs. The predicted model for LhKNOT was evaluated by Verify3D (<https://www.doe-mbi.ucla.edu/verify3d> ; accessed on 18 April 2018) [39], VoroMQA (<https://bioinformatics.lt/wtsam/voromqa> ; accessed on 18 April 2018) [40], and Prosa-web (<https://prosa.services.came.sbg.ac.at/prosa.php> ; accessed on 18 April 2018) [38]. Evaluating small peptides with standard evaluation tools can have limitations. Thus, scores from solved tertiary structures of NMR and crystallography-derived peptides of similar size and shape (PDB: 1Q3J, 1IVA, and 1DKC) were used to interpret the evaluation scores of our LhKNOT models. The top-ranked model was created using HHPred [37], a modeling program that uses the underlying model-building program, Modeller [86] using a restraint-based comparative modeling approach. In addition, the KNOTTIN database's KNOTER3D tool (<https://www.dsimb.inserm.fr/KNOTTIN/knoter3d.php> ; accessed on 15 May 2020) [36] was used to evaluate the model's tertiary structure, confirming LhKNOT's adoption of the Knottin fold. All the identified homologs with no known structures were modeled using the top template identified in HHPred [37].

4.4. CPC Clip Motif Identification

Superposition of LhKNOT's structural model with the known structures of a CPC clip containing knottin peptide from the cactus *Pereskia bleo* (PDB: 5XBD) [26] using the programs Superpose (<http://superpose.wishartlab.com> ; accessed on 12 September 2018) [87] and MultiProt (<http://bioinfo3d.cs.tau.ac.il/MultiProt> ; accessed on 12 September 2018) [88] suggested the presence of a heparin-binding CPC clip motif. To confirm the presence of the CPC clip in LhKNOT, docking analysis using the program, ClusPro 2.0 (which offers parameters specific to heparin as ligand; <https://cluspro.org> ; accessed on 2 March 2019) [89], and subsequent analysis of predicted docking scenarios in the visualization program, PyMOL (<https://pymol.org> ; accessed on 10 March 2019) [90] was performed. All potential docking scenarios were analyzed for identifying residues forming the putative CPC clip motif using the following steps. All polar interactions between LhKNOT and heparin were visualized in PyMOL and lists of participating residues were compiled. Predictions that contained the necessary Cation–Polar–Cation interactions were then further scrutinized for adherence to the CPC clip motif parameter by measuring distances between alpha carbons of participating residues as well as distances between center of mass of participating residues. PyMOL's "Measurements" Wizard tool was used to measure distances between α -carbons of the participating residues. To evaluate the distances between centers of mass, a user-generated script available in the pymolwiki website (Henschel, https://pymolwiki.org/index.php/Center_of_mass ; accessed on 15 March 2019) was used to generate pseudo-atoms at the calculated sidechain center of mass. PyMOL's APBS Electrostatics Plugin [91] was used to render the surface electrostatics for the modeled

LhKNOT. Similar steps were carried out to identify the CPC clip motif in each of the remaining knottin proteins in [Supplementary Table S5](#).

4.5. Association of LhKNOT with Other GAGs

To assess whether the CPC clip would facilitate binding with additional GAGs, further docking was performed with hyaluronic acid (PDB: 1HYA) and keratan sulfate (PDB: 1KES). In the case of hyaluronic acid, extraneous atoms (O, Na) were removed from the space surrounding the molecule. Keratan sulfate's model was used as is. Docking was performed using AutoDock (<https://autodock.scripps.edu> ; accessed on 7 April 2020) [92], and results were visualized in PyMOL. Results were then input into PDBSum (<http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html> ; accessed on 10 April 2020) [93] to generate Ligplots to confirm GAG-CPC clip binding.

4.6. Protein-Heparin Docking Analysis for Mutated CPC Clip in LhKNOT

To investigate the importance of the CPC clip in heparin binding, mutated models of LhKNOT were generated using PyMOL's mutagenesis tool, substituting the CPC clip residues with alanine. Models were created piecewise, first substituting just R14, then R1 and R14, and finally R14, R1, and S3. Heparin docking of all models was performed using ClusPro, and analysis was performed as described previously.

4.7. Protein-Heparin Docking Analysis in Structural Homologs and Drosomycin

To further evaluate the presence of the CPC clip motif in the knottin peptide family, the previous methods were further expanded to (a) structural homologs of LhKNOT (PDB: 5XBD, 1Q3J, 1IVA, and 1DKC) [25,26,28,30]; (b) *Drosophila* Drosomycin (PDB: 1MYN) [43], a peptide with known antifungal activity; (c) modeled homologous knottin proteins, and (d) representative knottins from different categories in the KNOTTIN database. Results were considered to adhere to the motif if measurements fell within 3 Å of the CPC clip parameters, to allow for some flexibility of protein loop conformations.

Acknowledgments

We are grateful to Brian Wey for help with *G. brasiliensis* gene predictions and S.A. Tasnim Ahmed for his help with formatting the figures.

Supplementary Materials

The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/pathogens12010143/s1>. Table S1. Identified homologs of LhKNOT using different sequence-based approaches. Table S2. Putative *L. heterotoma* homologs of LhKNOT. Sequences were identified via tblastn from three Lh14 transcriptome datasets (see Methods for accession numbers). Transcripts grouped by shaded cells encode identical knottin peptides. Transcript accession numbers shown in bold encode two distinct peptides. Results were retrieved from the Description Table available from NCBI tblastn results. Table S3. Putative *L. boulardi* knottin homologs of LhKNOT. Sequences were identified via tblastn from four *L. boulardi* transcriptome datasets (see Methods for accession numbers). Identical knottin peptides derived from different transcripts are grouped by color. Results were retrieved from the Description Table from NCBI tblastn results. Table S4. Putative knottin sequences from *Ganaspis* spp., homologous to LhKNOT. *G. hookeri* sequences were identified with tblastn, while *G. brasiliensis* sequences were identified with blastp. Results for *G. hookeri* and *G. brasiliensis* were retrieved from the Description Table available from NCBI search results. Shaded sequences encode identical peptides. Table S5. A list of amino acid residues in the CPC clips of examined known knottin, modeled knottin peptides, and Drosomycin. Distance measurements (Å) of alpha-carbon and sidechain center of mass (via pseudo-atom) of participating residues of identified CPC clip motifs are shown. Key: + Known offensive knottins (toxins), Φ in silico models, * non-knottin fold, ±alternate putative peptide, highlighted in green and yellow to demarcate peptides arising from the same transcript. Figure S1: (A) Prosa-web evaluation profile. Z-score of the model is represented as a black dot on a backdrop of Z-scores of known structures from the pdb (left) and visualization of the three-dimensional model LhKNOT with residues colored from blue to red in the order of increasing residue energy (right). (B) Verify3d evaluation profile. The plot shows the averaged and raw 1D-3D scores for LhKNOT and panel on top lists the percentage of residues that scored an average of 0.2 or above. A passing model has at least 80% of its residues scoring above 0.2. (C) VoroMQA evaluation profile. VoroMQA global scores in the context of high-quality X-ray structure scores (left) and plot of the local score and overall score for LhKNOT (right). Figure S2. Ligplot showing interactions of LhKNOT with keratan sulfate (PDB: 1KES) include the CPC clip residues (R1-S3-R14). Green dashed lines depict hydrogen bond interactions. Figure S3. Ligplot showing interactions of LhKNOT with hyaluronic acid (PDB: 1HYA) include the CPC clip residues (R1-S3-R14). Green dashed lines depict hydrogen bond interactions. Figure S4. (A) Ligplot showing interactions of wild type LhKNOT interacting with heparin via with the primary CPC clip (R1-S3-R14). (B) Ligplot showing interactions of mutant R14A LhKNOT interacting with heparin via a secondary CPC clip (R1-S3-R30). Green dashed lines depict hydrogen bond interactions. References [94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109,110,111] are cited in the Supplementary Materials.

[Click here for additional data file.](#) (1MB, zip)

Author Contributions

Conceptualization, S.S., S.G. and J.A.; methodology, J.A. and S.S.; validation, S.S., S.G., J.A., J.L. and L.V.C.; formal

analysis, J.A., J.L. and L.V.C.; investigation, J.A., J.L. and L.V.C.; resources, S.S. and S.G.; data curation, J.A., J.L. and L.V.C.; writing—original draft preparation, J.A.; writing—review and editing, S.S., S.G., J.A., J.L. and L.V.C.; visualization, J.A. and J.L.; supervision, S.S. and S.G.; project administration, S.S. and S.G. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The data presented in this study are available in the article and [Supplementary Materials](#).

Conflicts of Interest

The authors declare no conflict of interest.

Funding Statement

This research received no funding.

Footnotes

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

References

1. Penczykowski R.M., Laine A.L., Koskella B. Understanding the ecology and evolution of host-parasite interactions across scales. *Evol. Appl.* 2015;9:37–52. doi: 10.1111/eva.12294. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
2. Lambrechts L., Fellous S., Koella J.C. Coevolutionary interactions between host and parasite genotypes. *Trends Parasitol.* 2006;22:12–16. doi: 10.1016/j.pt.2005.11.008. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
3. Quicke D.L. *Parasitic Wasps*. Springer; Dordrecht, The Netherlands: 1997. [[Google Scholar](#)]
4. Poirié M., Colinet D., Gatti J. Insights into function and evolution of parasitoid wasp venoms. *Curr. Opin. Insect Sci.* 2014;6:52–60. doi: 10.1016/j.cois.2014.10.004. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
5. Kraaijeveld A.R., Godfray H.C.J. Evolution of host resistance and parasitoid counter-resistance. *Adv. Parasitol.* 2009;70:257–280. doi: 10.1016/S0065-308X(09)70010-7. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
6. Burke G.R., Strand M.R. Systematic analysis of a wasp parasitism arsenal. *Mol. Ecol.* 2014;23:890–901. doi: 10.1111/mec.12648. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
7. Yan Z., Fang Q., Wang L., Liu J., Zhu Y., Wang F., Li F., Werren J.H., Ye G. Insights into the venom composition and evolution of an endoparasitoid wasp by combining proteomic and transcriptomic analyses. *Sci. Rep.* 2016;6:19604. doi: 10.1038/srep19604. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
8. Lue C., Driskell A.C., Leips J., Buffington M.L. Review of the genus *Leptopilina* (Hymenoptera, Cynipoidea, Figitidae, Eucoilinae) from the Eastern United States, including three newly described species. *J. Hymenopt. Res.* 2016;53:35–76. doi: 10.3897/jhr.53.10369. [[DOI](#)] [[Google Scholar](#)]
9. Heavner M.E., Hudgins A.D., Rajwani R., Morales J., Govind S. Harnessing the natural *Drosophila*-parasitoid model for integrating insect immunity with functional venomomics. *Curr. Opin. Insect Sci.* 2014;6:61–67. doi: 10.1016/j.cois.2014.09.016. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
10. Lee M.J., Kalamarz M.E., Paddibhatla I., Small C., Rajwani R., Govind S. Virulence factors and strategies of *Leptopilina* spp.: Selective responses in *Drosophila* hosts. *Adv. Parasitol.* 2009;70:123–145. doi: 10.1016/S0065-308X(09)70005-3. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
11. Buchon N., Silverman N., Cherry S. Immunity in *Drosophila melanogaster*—From microbial recognition to whole-organism physiology. *Nat. Rev. Immunol.* 2014;14:796–810. doi: 10.1038/nri3763. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
12. Govind S. Innate immunity in *Drosophila*: Pathogens and pathways. *Insect Sci.* 2008;15:29–43. doi: 10.1111/j.1744-7917.2008.00185.x. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
13. Kounatidis I., Ligoxygakis P. *Drosophila* as a model system to unravel the layers of innate immunity to

infection. *Open Biol.* 2012;2:120075. doi: 10.1098/rsob.120075. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

14. Schlenke T.A., Morales J., Govind S., Clark A.G. Contrasting infection strategies in generalist and specialist wasp parasitoids of *Drosophila melanogaster*. *PLoS Pathog.* 2007;3:1486–1501. doi: 10.1371/journal.ppat.0030158. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

15. Qiu P., Pan P.C., Govind S. A role for the *Drosophila* Toll/Cactus pathway in larval hematopoiesis. *Development.* 1998;125:1909–1920. doi: 10.1242/dev.125.10.1909. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

16. Gueguen G., Kalamarz M.E., Ramroop J., Uribe J., Govind S. Polydn viral ankyrin proteins aid parasitic wasp survival by coordinate and selective inhibition of hematopoietic and immune NF-kappa B signaling in insect hosts. *PLoS Pathog.* 2013;9:e1003580. doi: 10.1371/journal.ppat.1003580. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

17. Louradour I., Sharma A., Morin-Poulard I., Letourneau M., Vincent A., Crozatier M., Vanzo N. Reactive oxygen species-dependent Toll/NF-kappaB activation in the *Drosophila* hematopoietic niche confers resistance to wasp parasitism. *eLife.* 2017;6:e25496. doi: 10.7554/eLife.25496. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

18. Sorrentino R.P., Melk J.P., Govind S. Genetic analysis of contributions of dorsal group and JAK-Stat92E pathway genes to larval hemocyte concentration and the egg encapsulation response in *Drosophila*. *Genetics.* 2004;166:1343–1356. doi: 10.1534/genetics.166.3.1343. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

19. Chiu H., Govind S. Natural infection of *D. melanogaster* by virulent parasitic wasps induces apoptotic depletion of hematopoietic precursors. *Cell Death Differ.* 2002;9:1379–1381. doi: 10.1038/sj.cdd.4401134. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

20. Rizki R.M., Rizki T.M. Parasitoid virus-like particles destroy *Drosophila* cellular immunity. *Proc. Natl. Acad. Sci. USA.* 1990;87:8388–8392. doi: 10.1073/pnas.87.21.8388. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

21. Heavner M.E., Ramroop J., Gueguen G., Ramrattan G., Dolios G., Scarpati M., Kwiat J., Bhattacharya S., Wang R., Singh S., et al. Novel Organelles with Elements of Bacterial and Eukaryotic Secretion Systems Weaponize Parasites of *Drosophila*. *Curr Biol.* 2017;27:2869–2877.e6. doi: 10.1016/j.cub.2017.08.019. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

22. Ferrarese R., Morales J., Fimiarz D., Webb B.A., Govind S. A supracellular system of actin-lined canals controls biogenesis and release of virulence factors in parasitoid venom glands. *J. Exp. Biol.* 2009;212:2261–2268. doi: 10.1242/jeb.025718. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

23. Chiu H., Morales J., Govind S. Identification and immuno-electron microscopy localization of p40, a protein component of immunosuppressive virus-like particles from *Leptopilina heterotoma*, a virulent parasitoid wasp of *Drosophila*. *J. Gen. Virol.* 2006;87:461–470. doi: 10.1099/vir.0.81474-0. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
24. Gracy J., Chiche L. Structure and modeling of knottins, a promising molecular scaffold for drug discovery. *Curr. Pharm. Des.* 2011;17:4337–4350. doi: 10.2174/138161211798999339. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
25. Gao G.H., Liu W., Dai J.X., Wang J.F., Hu Z., Zhang Y., Wang D.C. Solution structure of PAFP-S: A new knottin-type antifungal peptide from the seeds of *Phytolacca americana*. *Biochemistry.* 2001;40:10973–10978. doi: 10.1021/bi010167k. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
26. Loo S., Kam A., Xiao T., Tam J.P. Bleogens, Cactus-Derived Anti-Candida Cysteine-Rich Peptides with Three Different Precursor Arrangements. *Front. Plant Sci.* 2017;8:2162. doi: 10.3389/fpls.2017.02162. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
27. Becker S., Terlau H. Toxins from cone snails: Properties, applications and biotechnological production. *Appl. Microbiol. Biotechnol.* 2008;79:1–9. doi: 10.1007/s00253-008-1385-6. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
28. Reily M.D., Holub K.E., Gray W.R., Norris T.M., Adams M.E. Structure-activity relationships for P-type calcium channel-selective omega-agatoxins. *Nat. Struct. Biol.* 1994;1:853–856. doi: 10.1038/nsb1294-853. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
29. Nguyen K.N.T., Nguyen G.K.T., Nguyen P.Q.T., Ang K.H., Dedon P.C., Tam J.P. Immunostimulating and Gram-negative-specific antibacterial cyclotides from the butterfly pea (*Clitoria ternatea*) *FEBS J.* 2016;283:2067–2090. doi: 10.1111/febs.13720. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
30. Barbault F., Landon C., Guenneugues M., Meyer J.P., Schott V., Dimarcq J.L., Vovelle F. Solution structure of Alo-3: A new knottin-type antifungal peptide from the insect *Acrocinus longimanus*. *Biochemistry.* 2003;42:14434–14442. doi: 10.1021/bi035400o. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
31. Cox N., Kintzing J.R., Smith M., Grant G.A., Cochran J.R. Integrin-Targeting Knottin Peptide-Drug Conjugates Are Potent Inhibitors of Tumor Cell Proliferation. *Angew. Chem. Int. Ed. Engl.* 2016;55:9894–9897. doi: 10.1002/anie.201603488. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
32. Zhang C.R., Zhang S., Xia J., Li F.F., Xia W.Q., Liu S.S., Wang X.W. The immune strategy and stress response of the Mediterranean species of the *Bemisia tabaci* complex to an orally delivered bacterial pathogen. *PLoS ONE.* 2014;9:e94477. doi: 10.1371/journal.pone.0094477. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

33. Shao F., Hu Z., Xiong Y.M., Huang Q.Z., Wang C.G., Zhu R.H., Wang D.C. A new antifungal peptide from the seeds of *Phytolacca americana*: Characterization, amino acid sequence and cDNA cloning. *Biochim. Biophys. Acta*. 1999;1430:262–268. doi: 10.1016/S0167-4838(99)00013-8. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
34. Fletcher J.I., Smith R., O'Donoghue S.I., Nilges M., Connor M., Howden M.E., Christie M.J., King G.F. The structure of a novel insecticidal neurotoxin, omega-atracotoxin-HV1, from the venom of an Australian funnel web spider. *Nat. Struct. Biol.* 1997;4:559–566. doi: 10.1038/nsb0797-559. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
35. Casale J., Crane J.S. Anonymous StatPearls. StatPearls Publishing LLC.; Treasure Island, FL, USA: 2022. Biochemistry, Glycosaminoglycans. [[Google Scholar](#)]
36. Postic G., Gracy J., Périn C., Chiche L., Gelly J. KNOTTIN: The database of inhibitor cystine knot scaffold after 10 years, toward a systematic structure modeling. *Nucleic Acids Res.* 2018;46:D454–D458. doi: 10.1093/nar/gkx1084. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
37. Soding J., Biegert A., Lupas A.N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 2005;33((Suppl. 2)):244. doi: 10.1093/nar/gki408. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
38. Wiederstein M., Sippl M.J. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 2007;35:407. doi: 10.1093/nar/gkm290. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
39. Eisenberg D., Luthy R., Bowie J.U. VERIFY3D: Assessment of protein models with three-dimensional profiles. *Methods Enzymol.* 1997;277:396–404. doi: 10.1016/S0076-6879(97)77022-8. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
40. Olechnovič K., Venclovas Č. VoroMQA: Assessment of protein structure quality using interatomic contact areas. *Proteins*. 2017;85:1131–1145. doi: 10.1002/prot.25278. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
41. Pulido D., Rebolledo-Rios R., Valle J., Andreu D., Boix E., Torrent M. Structural similarities in the CPC clip motif explain peptide-binding promiscuity between glycosaminoglycans and lipopolysaccharides. *J. R. Soc. Interface*. 2017;14:20170423. doi: 10.1098/rsif.2017.0423. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
42. Torrent M., Nogues M.V., Andreu D., Boix E. The “CPC clip motif”: A conserved structural signature for heparin-binding proteins. *PLoS ONE*. 2012;7:e42692. doi: 10.1371/journal.pone.0042692. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

43. Landon C., Sodano P., Hetru C., Hoffmann J., Ptak M. Solution structure of drosomycin, the first inducible antifungal protein from insects. *Protein Sci.* 1997;6:1878–1884. doi: 10.1002/pro.5560060908. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
44. Cornet B., Bonmatin J.M., Hetru C., Hoffmann J.A., Ptak M., Vovelle F. Refined three-dimensional solution structure of insect defensin A. *Structure.* 1995;3:435–448. doi: 10.1016/S0969-2126(01)00177-0. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
45. Illergard K., Ardell D.H., Elofsson A. Structure is three to ten times more conserved than sequence—A study of structural response in protein cores. *Proteins.* 2009;77:499–508. doi: 10.1002/prot.22458. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
46. Wywiał E., Singh S.M. Identification and structural characterization of FYVE domain-containing proteins of *Arabidopsis thaliana*. *BMC Plant Biol.* 2010;10:157. doi: 10.1186/1471-2229-10-157. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
47. Martin C.L., Singh S.M. Identification, modeling, and characterization studies of *Tetrahymena thermophila* myosin FERM domains suggests a conserved core fold but functional differences. *Cytoskeleton.* 2015;72:585–596. doi: 10.1002/cm.21261. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
48. Wiech E.M., Cheng H.P., Singh S.M. Molecular modeling and computational analyses suggests that the *Sinorhizobium meliloti* periplasmic regulator protein ExoR adopts a superhelical fold and is controlled by a unique mechanism of proteolysis. *Protein Sci.* 2015;24:319–327. doi: 10.1002/pro.2616. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
49. Scarpati M., Qi Y., Govind S., Singh S. A combined computational strategy of sequence and structural analysis predicts the existence of a functional eicosanoid pathway in *Drosophila melanogaster*. *PLoS ONE.* 2019;14:e0211897. doi: 10.1371/journal.pone.0211897. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
50. Barreto C., Silva A., Wiech E., Lopez A., San A., Singh S. Proteomic Tools for the Analysis of Cytoskeleton Proteins. *Methods Mol. Biol.* 2022;2364:363–425. doi: 10.1007/978-1-0716-1661-1_19. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
51. Raskob G.E., Carter C.J., Hull R.D. Heparin therapy for venous thrombosis and pulmonary embolism. *Blood Rev.* 1988;2:251–258. doi: 10.1016/0268-960X(88)90014-8. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
52. Zhang L. Glycosaminoglycan (GAG) biosynthesis and GAG-binding proteins. *Prog. Mol. Biol. Transl. Sci.* 2010;93:1–17. doi: 10.1016/S1877-1173(10)93001-9. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
53. Andersson E., Rydengard V., Sonesson A., Morgelin M., Björck L., Schmidtchen A. Antimicrobial

activities of heparin-binding peptides. *Eur. J. Biochem.* 2004;271:1219–1226. doi: 10.1111/j.1432-1033.2004.04035.x. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

54. Iannuzzi C., Irace G., Sirangelo I. The effect of glycosaminoglycans (GAGs) on amyloid aggregation and toxicity. *Molecules.* 2015;20:2510–2528. doi: 10.3390/molecules20022510. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

55. Zhang Z., Zhu S. Functional role of charged residues in drosomycin, a *Drosophila* antifungal peptide. *Dev. Comp. Immunol.* 2010;34:953–958. doi: 10.1016/j.dci.2010.04.003. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

56. Govind S., Nehm R.H. Innate immunity in fruit flies: A textbook example of genomic recycling. *PLoS Biol.* 2004;2:E276. doi: 10.1371/journal.pbio.0020276. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

57. Araki M., Kurihara M., Kinoshita S., Awane R., Sato T., Ohkawa Y., Inoue Y.H. Anti-tumour effects of antimicrobial peptides, components of the innate immune system, against haematopoietic tumours in *Drosophila mxc* mutants. *Dis. Model Mech.* 2019;12:dmm037721. doi: 10.1242/dmm.037721. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

58. Ramroop J.R., Heavner M.E., Razzak Z.H., Govind S. A parasitoid wasp of *Drosophila* employs preemptive and reactive strategies to deplete its host's blood cells. *PLoS Pathog.* 2021;17:e1009615. doi: 10.1371/journal.ppat.1009615. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

59. Zhu F., Li D., Chen K. Structures and functions of invertebrate glycosylation. *Open Biol.* 2019;9:180232. doi: 10.1098/rsob.180232. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

60. Iozzo R.V., Schaefer L. Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. *Matrix Biol.* 2015;42:11–55. doi: 10.1016/j.matbio.2015.02.003. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

61. Alexander C., Rietschel E.T. Bacterial lipopolysaccharides and innate immunity. *J. Endotoxin Res.* 2001;7:167–202. doi: 10.1179/096805101101532675. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

62. Baron M.J., Wong S.L., Nybakken K., Carey V.J., Madoff L.C. Host glycosaminoglycan confers susceptibility to bacterial infection in *Drosophila melanogaster*. *Infect Immun.* 2009;77:860–866. doi: 10.1128/IAI.00995-08. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

63. Imler J., Bulet P. Antimicrobial peptides in *Drosophila*: Structures, activities and gene regulation. *Chem Immunol Allerg.* 2005;86:1–21. doi: 10.1159/000086648. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

64. Hwang B., Hwang J., Lee J., Lee D.G. Antifungal properties and mode of action of psacothecin, a novel

knottin-type peptide derived from *Psacotha hilaris*. *Biochem. Biophys. Res. Commun.* 2010;400:352–357. doi: 10.1016/j.bbrc.2010.08.063. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

65. De Guillen K., Lorrain C., Tsan P., Barthe P., Petre B., Saveleva N., Rouhier N., Duplessis S., Padilla A., Hecker A. Structural genomics applied to the rust fungus *Melampsora larici-populina* reveals two candidate effector proteins adopting cystine knot and NTF2-like protein folds. *Sci. Rep.* 2019;9:18084–18089. doi: 10.1038/s41598-019-53816-9. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

66. NCBI Resource Coordinators Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 2018;46:D8–D13. doi: 10.1093/nar/gkx1095. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

67. Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., Bourne P.E. The Protein Data Bank. *Nucleic Acids Res.* 2000;28:235–242. doi: 10.1093/nar/28.1.235. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

68. Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. Basic local alignment search tool. *J. Mol. Biol.* 1990;215:403–410. doi: 10.1016/S0022-2836(05)80360-2. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

69. Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* 1997;25:3389–3402. doi: 10.1093/nar/25.17.3389. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

70. Madeira F., Pearce M., Tivey A.R., Basutkar P., Lee J., Edbali O., Madhusoodanan N., Kolesnikov A., Lopez R. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* 2022;50:276. doi: 10.1093/nar/gkac240. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

71. Finn R.D., Clements J., Eddy S.R. HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Res.* 2011;39:29. doi: 10.1093/nar/gkr367. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

72. Goecks J., Mortimer N.T., Mobley J.A., Bowersock G.J., Taylor J., Schlenke T.A. Integrative approach reveals composition of endoparasitoid wasp venoms. *PLoS ONE.* 2013;8:e64125. doi: 10.1371/journal.pone.0064125. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

73. Varaldi J., Lepetit D. Deciphering the behaviour manipulation imposed by a virus on its parasitoid host: Insights from a dual transcriptomic approach. *Parasitology.* 2018;145:1979–1989. doi: 10.1017/S0031182018000835. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

74. Mortimer N.T., Goecks J., Kacsoh B.Z., Mobley J.A., Bowersock G.J., Taylor J., Schlenke T.A. Parasitoid wasp venom SERCA regulates *Drosophila* calcium levels and inhibits cellular immunity. *Proc. Natl. Acad.*

Sci. USA. 2013;110:9427–9432. doi: 10.1073/pnas.1222351110. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

75. Gasteiger E., Gattiker A., Hoogland C., Ivanyi I., Appel R.D., Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 2003;31:3784–3788. doi: 10.1093/nar/gkg563. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

76. Keller O., Kollmar M., Stanke M., Waack S. A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics.* 2011;27:757–763. doi: 10.1093/bioinformatics/btr010. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

77. Sievers F., Wilm A., Dineen D., Gibson T.J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Söding J., et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 2011;7:539. doi: 10.1038/msb.2011.75. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

78. Poirot O., O'Toole E., Notredame C. Tcoffee@ igs: A web server for computing, evaluating and combining multiple sequence alignments. *Nucleic Acids Res.* 2003;31:3503–3506. doi: 10.1093/nar/gkg522. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

79. Gouet P., Courcelle E., Stuart D.I., Metoz F. ESPript: Analysis of multiple sequence alignments in PostScript. *Bioinformatics.* 1999;15:305–308. doi: 10.1093/bioinformatics/15.4.305. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

80. Marchler-Bauer A., Lu S., Anderson J.B., Chitsaz F., Derbyshire M.K., DeWeese-Scott C., Fong J.H., Geer L.Y., Geer R.C., Gonzales N.R., et al. CDD: A Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* 2011;39:225. doi: 10.1093/nar/gkq1189. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

81. Finn R.D., Bateman A., Clements J., Coggill P., Eberhardt R.Y., Eddy S.R., Heger A., Hetherington K., Holm L., Mistry J., et al. Pfam: The protein families database. *Nucleic Acids Res.* 2014;42:222. doi: 10.1093/nar/gkt1223. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

82. Almagro Armenteros J.J., Tsirigos K.D., Sønderby C.K., Petersen T.N., Winther O., Brunak S., von Heijne G., Nielsen H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 2019;37:420–423. doi: 10.1038/s41587-019-0036-z. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

83. Kall L., Krogh A., Sonnhammer E.L. Advantages of combined transmembrane topology and signal peptide prediction—The Phobius web server. *Nucleic Acids Res.* 2007;35:429. doi: 10.1093/nar/gkm256. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

84. Simossis V.A., Heringa J. The influence of gapped positions in multiple sequence alignments on secondary structure prediction methods. *Comput. Biol. Chem.* 2004;28:351–366. doi: 10.1016/j.compbiolchem.2004.09.005. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
85. Roy A., Kucukural A., Zhang Y. I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* 2010;5:725–738. doi: 10.1038/nprot.2010.5. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
86. Webb B., Sali A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinform.* 2014;47:5–6. doi: 10.1002/0471250953.bi0506s47. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
87. Maiti R., Van Domselaar G.H., Zhang H., Wishart D.S. SuperPose: A simple server for sophisticated structural superposition. *Nucleic Acids Res.* 2004;32:590. doi: 10.1093/nar/gkh477. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
88. Shatsky M., Nussinov R., Wolfson H.J. A method for simultaneous alignment of multiple protein structures. *Proteins.* 2004;56:143–156. doi: 10.1002/prot.10628. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
89. Vajda S., Yueh C., Beglov D., Bohnuud T., Mottarella S.E., Xia B., Hall D.R., Kozakov D. New additions to the ClusPro server motivated by CAPRI. *Proteins.* 2017;85:435–444. doi: 10.1002/prot.25219. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
90. Schrodinger LLC . The PyMOL Molecular Graphics System. Schrodinger, LLC; New York, NY, USA: 2015. Version 1.8. in press. [[Google Scholar](#)]
91. Baker N.A., Sept D., Joseph S., Holst M.J., McCammon J.A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci USA.* 2001;98:10037–10041. doi: 10.1073/pnas.181342398. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
92. Morris G.M., Goodsell D.S., Halliday R.S., Huey R., Hart W.E., Belew R.K., Olson A.J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* 1998;19:1639–1662. doi: 10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B. [[DOI](#)] [[Google Scholar](#)]
93. Laskowski R.A., Jabłońska J., Pravda L., Vařeková R.S., Thornton J.M. PDBsum: Structural summaries of PDB entries. *Protein Sci.* 2018;27:129–134. doi: 10.1002/pro.3289. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
94. Lippens G., Najib J., Wodak S.J., Tartar A. NMR sequential assignments and solution structure of chlorotoxin, a small scorpion toxin that blocks chloride channels. *Biochemistry.* 1995;34:13–21. doi: 10.1021/bi00001a003. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

95. Takahashi H., Kim J.I., Min H.J., Sato K., Swartz K.J., Shimada I. Solution structure of hanatoxin1, a gating modifier of voltage-dependent K(+) channels: Common surface features of gating modifier toxins. *J. Mol. Biol.* 2000;297:771–780. doi: 10.1006/jmbi.2000.3609. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
96. Landon C., Pajon A., Vovelle F., Sodano P. The active site of drosomycin, a small insect antifungal protein, delineated by comparison with the modeled structure of Rs-AFP2, a plant antifungal protein. *J. Pept. Res.* 2000;56:231–238. doi: 10.1034/j.1399-3011.2000.00757.x. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
97. Hill J.M., Alewood P.F., Craik D.J. Solution structure of the sodium channel antagonist conotoxin GS: A new molecular caliper for probing sodium channel geometry. *Structure.* 1997;5:571–583. doi: 10.1016/S0969-2126(97)00212-8. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
98. Nielsen K.J., Thomas L., Lewis R.J., Alewood P.F., Craik D.J. A consensus structure for omega-conotoxins with different selectivities for voltage-sensitive calcium channel subtypes: Comparison of MVIIA, SVIB and SNX-202. *J. Mol. Biol.* 1996;263:297–310. doi: 10.1006/jmbi.1996.0576. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
99. Scanlon M.J., Naranjo D., Thomas L., Alewood P.F., Lewis R.J., Craik D.J. Solution structure and proposed binding mechanism of a novel potassium channel toxin kappa-conotoxin PVIIA. *Structure.* 1997;5:1585–1597. doi: 10.1016/S0969-2126(97)00307-9. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
100. Fujitani N., Kawabata S., Osaki T., Kumaki Y., Demura M., Nitta K. Structure of the antimicrobial peptide tachystatin A. *J. Biol. Chem.* 2002;277:23651–23657. doi: 10.1074/jbc.M111120200. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
101. Fujitani N., Kouno T., Nakahara T., Takaya K., Osaki T., Kawabata S. The solution structure of horseshoe crab antimicrobial peptide tachystatin B with an inhibitory cystine-knot motif. *J. Pept. Sci.* 2007;13:269–279. doi: 10.1002/psc.846. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
102. Bernard C., Corzo G., Mosbah A., Nakajima T., Darbon H. Solution structure of Ptu1, a toxin from the assassin bug *Peirates turpis* that blocks the voltage-sensitive calcium channel N-type. *Biochemistry.* 2001;40:12795–12800. doi: 10.1021/bi015537j. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
103. Bernard C., Corzo G., Adachi-Akahane S., Foures G., Kanemaru K., Furukawa Y. Solution structure of ADO1, a toxin extracted from the saliva of the assassin bug, *Agriosphodrus dohrni*. *Proteins.* 2004;54:195–205. doi: 10.1002/prot.10513. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
104. Martins J.C., Enassar M., Willem R., Wieruzeski J.M., Lippens G., Wodak S.J. Solution structure of the main alpha-amylase inhibitor from amaranth seeds. *Eur. J. Biochem.* 2001;268:2379–2389. doi: 10.1046/j.1432-1327.2001.02118.x. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

105. Janssen B.J.C., Schirra H.J., Lay F.T., Anderson M.A., Craik D.J. Structure of *Petunia hybrida* defensin 1, a novel plant defensin with five disulfide bonds. *Biochemistry*. 2003;42:8214–8222. doi: 10.1021/bi034379o. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
106. Frenal K., Xu C., Wolff N., Wecker K., Gurrola G.B., Zhu S., Chi C.-W., Possani L.D., Tytgat J., Delepierre M. Exploring structural features of the interaction between the scorpion toxin CnErg1 and ERG K⁺ channels. *Proteins*. 2004;56:367–375. doi: 10.1002/prot.20102. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
107. Lomize A.L., Maiorov V.N., Arsen'ev A.S. Determination of the spatial structure of insectotoxin 15A from *Buthus erpeus* by (1)H-NMR spectroscopy data. *Bioorg Khim*. 1991;17:1613–1632. [[PubMed](#)] [[Google Scholar](#)]
108. Pimentel C., Choi S., Chagot B., Guette C., Camadro J., Darbon H. Solution structure of PcFK1, a spider peptide active against *Plasmodium falciparum*. *Protein Sci*. 2006;15:628–634. doi: 10.1110/ps.051860606. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
109. Koltay A., Daly N.L., Gustafson K.R., Craik D.J. Structure of Circulin B and Implications for Antimicrobial Activity of the Cyclotides. *Int. J. Pept. Res. Ther*. 2005;11:99–106. doi: 10.1007/s10989-004-1722-2. [[DOI](#)] [[Google Scholar](#)]
110. Craik D.J., Daly N.L., Bond T., Waine C. Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J. Mol. Biol*. 1999;294:1327–1336. doi: 10.1006/jmbi.1999.3383. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]
111. Li H., Bowling J.J., Fronczek F.R., Hong J., Jabba S.V., Murray T.F., Ha N.-C., Hamann M.T., Jung J.H. Asteropsin A: An unusual cystine-crosslinked peptide from porifera enhances neuronal Ca²⁺ influx. *Biochim Biophys Acta*. 2013;1830:2591–2599. doi: 10.1016/j.bbagen.2012.11.015. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

Associated Data

This section collects any data citations, data availability statements, or supplementary materials included in this article.

Supplementary Materials

[Click here for additional data file.](#) (1MB, zip)

Data Availability Statement

The data presented in this study are available in the article and [Supplementary Materials](#).

Articles from Pathogens are provided here courtesy of **Multidisciplinary Digital Publishing Institute (MDPI)**